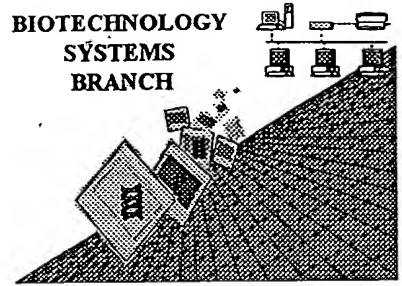


McElwain

RAW SEQUENCE LISTING **ERROR REPORT**

BIOTECHNOLOGY
SYSTEMS
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following CRF diskette:

Application Serial Number: 09/049,304
Art Unit / Team No. : 1649
Date Processed by STIC: 1/31/2000

THE ATTACHED PRINTOUT EXPLAINS THE ERRORS DETECTED.

PLEASE BE SURE TO FORWARD THIS INFORMATION TO THE APPLICANTS BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANTS ALONG WITH A NOTICE TO COMPLY or,**
- 2) CALLING APPLICANTS AND FAXING THEM A COPY OF THE PRINTOUT WITH A NOTICE TO COMPLY**

THIS WILL INSURE THAT THE NEXT SUBMISSION RECEIVED FROM THEM WILL BE ERROR FREE.

IF YOU HAVE ANY FURTHER QUESTIONS, PLEASE CALL:

MARK SPENCER 703-308-4212

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/049,304

DATE: 02/02/2000
TIME: 22:09:34

INPUT SET: S34610.raw

This Raw Listing contains the General Information Section and those Sequences containing ERRORS.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: EPELBAUM, SABINE URSULA
FALCO, SAVERIO CARL
MCDEVITT, RAYMOND ERVIN, III

(ii) TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR
INCREASING THE LYSINE CONTENT OF
THE SEEDS OF PLANTS

(iii) NUMBER OF SEQUENCES: 132

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
(B) STREET: 1007 MARKET STREET
(C) CITY: WILMINGTON
(D) STATE: DELAWARE
(E) COUNTRY: U.S.A.
(F) ZIP: 19898

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: DISKETTE, 3.50 INCH
(B) COMPUTER: IBM PC COMPATIBLE
(C) OPERATING SYSTEM: MICROSOFT OFFICE 97
(D) SOFTWARE: MICROSOFT WINDOWS 95

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/824,627
(B) FILING DATE: MARCH 27, 1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: CHRISTENBURY, LYNNE M.
(B) REGISTRATION NUMBER: 30,971
(C) REFERENCE/DOCKET NUMBER: BB-1037-F

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 302-992-5481
(B) TELEFAX: 302-892-7949

Does Not Comply
Corrected Diskette Needed

INPUT SET: S34610.raw

46 (C) TELEX: 835420
47
48

ERRORED SEQUENCES FOLLOW:

237 (2) INFORMATION FOR SEQ ID NO:6:
238
239 (i) SEQUENCE CHARACTERISTICS: (see p. 3)
240 (A) LENGTH: 917 base pairs
241 (B) TYPE: nucleic acid
242 (C) STRANDEDNESS: single
243 (D) TOPOLOGY: linear
244
245 (ii) MOLECULE TYPE: DNA (genomic)
246
247 (ix) FEATURE:
248 (A) NAME/KEY: CDS
249 (B) LOCATION: 3..911
250
251 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
252
253 CC ATG GCT ACA GGT TTA ACA GCT AAG ACC GGA GTA GAG CAC TTC GGC 47
254 Met Ala Thr Gly Leu Thr Ala Lys Thr Gly Val Glu His Phe Gly
255 1 5 10 15
256
257 ACC GTT GGA GTA GCA ATG GTT ACT CCA TTC ACG GAA TCC GGA GAC ATC 95
258 Thr Val Gly Val Ala Met Val Thr Pro Phe Thr Glu Ser Gly Asp Ile
259 20 25 30
260
261 GAT ATC GCT GCT GGC CGC GAA GTC GCG GCT TAT TTG GTT GAT AAG GGC 143
262 Asp Ile Ala Ala Gly Arg Glu Val Ala Ala Tyr Leu Val Asp Lys Gly
263 35 40 45
264
265 TTG GAT TCT TTG GTT CTC GCG GGC ACC ACT GGT GAA TCC CCA ACG ACA 191
266 Leu Asp Ser Leu Val Leu Ala Gly Thr Thr Gly Glu Ser Pro Thr Thr
267 50 55 60
268
269 ACC GCC GCT GAA AAA CTA GAA CTG CTC AAG GCC GTT CGT GAG GAA GTT 239
270 Thr Ala Ala Glu Lys Leu Glu Leu Leu Lys Ala Val Arg Glu Glu Val
271 65 70 75
272
273 GGG GAT CGG GCG AAG CTC ATC GCC GGT GTC GGA ACC AAC AAC ACG CGG 287
274 Gly Asp Arg Ala Lys Leu Ile Ala Gly Val Gly Thr Asn Asn Thr Arg
275 80 85 90 95
276
277 ACA TCT GTG GAA CTT GCG GAA GCT GCT GCT TCT GCT GGC GCA GAC GGC 335
278 Thr Ser Val Glu Leu Ala Glu Ala Ala Ser Ala Gly Ala Asp Gly
279 100 105 110
280

RAW SEQUENCE LISTING PATENT APPLICATION US/09/049,304

DATE: 02/02/2000
TIME: 22:09:35

INPUT SET: S34610.raw

```

281  CTT TTA GTT GTA ACT CCT TAT TAC TCC AAG CCG AGC CAA GAG GGA TTG  383
282  Leu Leu Val Val Thr Pro Tyr Tyr Ser Lys Pro Ser Gln Glu Gly Leu
283              115                      120                      125
284
285  CTG GCG CAC TTC GGT GCA ATT GCT GCA GCA ACA GAG GTT CCA ATT TGT  431
286  Leu Ala His Phe Gly Ala Ile Ala Ala Ala Thr Glu Val Pro Ile Cys
287              130                      135                      140
288
289  CTC TAT GAC ATT CCT GGT CGG TCA GGT ATT CCA ATT GAG TCT GAT ACC  479
290  Leu Tyr Asp Ile Pro Gly Arg Ser Gly Ile Pro Ile Glu Ser Asp Thr
291              145                      150                      155
292
293  ATG AGA CGC CTG AGT GAA TTA CCT ACG ATT TTG GCG GTC AAG GAC GCC  527
294  Met Arg Arg Leu Ser Glu Leu Pro Thr Ile Leu Ala Val Lys Asp Ala
295  160                      165                      170                      175
296
297  AAG GGT GAC CTC GTT GCA GCC ACG TCA TTG ATC AAA GAA ACG GGA CTT  575
298  Lys Gly Asp Leu Val Ala Ala Thr Ser Leu Ile Lys Glu Thr Gly Leu
299              180                      185                      190
300
301  GCC TGG TAT TCA GGC GAT GAC CCA CTA AAC CTT GTT TGG CTT GCT TTG  623
302  Ala Trp Tyr Ser Gly Asp Asp Pro Leu Asn Leu Val Trp Leu Ala Leu
303              195                      200                      205
304
305  GGC GGA TCA GGT TTC ATT TCC GTA ATT GGA CAT GCA GCC CCC ACA GCA  671
306  Gly Gly Ser Gly Phe Ile Ser Val Ile Gly His Ala Ala Pro Thr Ala
307              210                      215                      220
308
309  TTA CGT GAG TTG TAC ACA AGC TTC GAG GAA GGC GAC CTC GTC CGT GCG  719
310  Leu Arg Glu Leu Tyr Thr Ser Phe Glu Glu Gly Asp Leu Val Arg Ala
311              225                      230                      235
312
313  CGG GAA ATC AAC GCC AAA CTA TCA CCG CTG GTA GCT GCC CAA GGT CGC  767
314  Arg Glu Ile Asn Ala Lys Leu Ser Pro Leu Val Ala Ala Gln Gly Arg
315  240                      245                      250                      255
316
317  TTG GGT GGA GTC AGC TTG GCA AAA GCT GCT CTG CGT CTG CAG GGC ATC  815
318  Leu Gly Gly Val Ser Leu Ala Lys Ala Ala Leu Arg Leu Gln Gly Ile
319              260                      265                      270
320
321  AAC GTA GGA GAT CCT CGA CTT CCA ATT ATG GCT CCA AAT GAG CAG GAA  863
322  Asn Val Gly Asp Pro Arg Leu Pro Ile Met Ala Pro Asn Glu Gln Glu
323              275                      280                      285
324
--> 325  CTT GAG GCT CTC CGA GAA GAC ATG AAA AAA GCT GGA GTT CTA TAA TGAGAATTC  918
326  Leu Glu Ala Leu Arg Glu Asp Met Lys Lys Ala Gly Val Leu *
327              290                      295                      300
328

```

918

917

1982 (2) INFORMATION FOR SEQ ID NO:76:

1983

1984 (i) SEQUENCE CHARACTERISTICS:

1985 (A) LENGTH: 175 base pairs

(next page)

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/049,304DATE: 02/02/2000
TIME: 22:09:35

INPUT SET: S34610.raw

1986 (B) TYPE: nucleic acid
1987 (C) STRANDEDNESS: double
1988 (D) TOPOLOGY: linear
1989
1990 (ii) MOLECULE TYPE: DNA (genomic)
1991
1992 (vi) ORIGINAL SOURCE:
1993 (B) STRAIN: E. coli
1994 (G) CELL TYPE: DH5 alpha
1995
1996 (vii) IMMEDIATE SOURCE:
1997 (B) CLONE: 5-1
1998
1999 (ix) FEATURE:
2000 (A) NAME/KEY: CDS
2001 (B) LOCATION: 2..172
2002 (D) OTHER INFORMATION: /function= "synthetic
2003 storage protein
2004 /product= "protein"
2005 /gene= "ssp"
2006 /standard_name=
2007 "5.5.5.7.7.7.5"
2008
2009 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:
2010
2011 C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG 46
2012 Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met
2013 1 5 10 15
2014
2015 GAG GAG AAG ATG AAG GCG ATG GAG GAA AAG CTG AAA GCG ATG GAG GAG 94
2016 Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu
2017 20 25 30
2018
2019 AAA CTC AAG GCT ATG GAA GAA AAG CTT AAA GCG ATG GAG GAG AAA CTG 142
2020 Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu
2021 35 40 45
2022
--> 2023 AAG GCC ATG GAA GAG AAG ATG AAG GCG TGATAG
2024 Lys Ala Met Glu Glu Lys Met Lys Ala
2025 50 55
2026

(179) 175

3026 (2) INFORMATION FOR SEQ ID NO:111:
3027
3028 (i) SEQUENCE CHARACTERISTICS:
--> 3029 (A) LENGTH: 3194 base pairs 3195 shown (p.7)
3030 (B) TYPE: nucleic acid
3031 (C) STRANDEDNESS: double
3032 (D) TOPOLOGY: linear
3033
3034 (ii) MOLECULE TYPE: DNA (genomic)
3035
3036 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

DATE: 02/02/2000
TIME: 22:09:36

INPUT SET: S34610.raw

3037								60
3038	ATGAATTCAA	ATGGCCATGA	GGAGGAGAAG	AAGTTGGGGA	ATGGAGTTGT	GGGGATTCTA		60
3039								
3040	TCTGAAACAG	TTAACAAATG	GGAGAGACGA	ACACCATTGA	CGCCATCGCA	TTGCGCTCGC		120
3041								
3042	CTTTTACACG	GTGGGAAAGA	CAGAACCGGC	ATTTCCCGCA	TTGTGGTTCA	GCCATCTGCT		180
3043								
3044	AAGCGTATCC	ATCATGATGC	CTTGTATGAA	CATGTTGGGT	GTGAAATTTT	TGATGATTTG		240
3045								
3046	TCTGATTGTG	GGCTTATACT	TGGAATCAAA	CAACCTGAGC	TAGAAATGAT	TCTTCCAGAG		300
3047								
3048	AGAGCATACG	CTTTCCTTTC	ACATACTCAT	AAGGCACAGA	AAGAGAACAT	GCCTTTGTTG		360
3049								
3050	GATAAAATTC	TTTCTGAGAG	AGTGACTTTG	TGTGATTATG	AGCTCATTGT	TGGGGATCAT		420
3051								
3052	GGGAAACGAT	TATTGGCGTT	TGGTAAATAT	GCAGGCAGAG	CTGGTCTTGT	TGACTTCTTA		480
3053								
3054	CACGGACTTG	GACAGCGATA	TCTAAGTCTA	GGATACTCAA	CACCTTTCCT	CTCGCTCGGT		540
3055								
3056	GCATCGTATA	TGTATTCCTC	ATTGGCTGCT	GCAAAAGCCG	CTGTAATTTT	TGTTGGTGAA		600
3057								
3058	GAAATTGCAA	GCCAGGGACT	GCCATTAGGA	ATCTGCCCTC	TTGTATTTGT	CTTACCCGGA		660
3059								
3060	ACAGGAAATG	TTTCTCTGGG	GGCGCAAGAA	ATTTTCAAGC	TTCTTCCTCA	CACTTTTGTT		720
3061								
3062	GAACCAAGCA	AACTTCCTGA	ACTATTTGTA	AAA				

Agave black section

[illegible]

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/049,304DATE: 02/02/2000
TIME: 22:09:36

INPUT SET: S34610.raw

3090	GCTGAATCGT	ACTCAGAACT	TGAAGTTGGT	GCGGATGATA	AGAGAGTATT	GGATCAAATC	1620
3091							
3092	ATTGATTCAT	TAACTCGGTT	AGCTAATCCA	AATGAAGATT	ATATATCCCC	ACATAGAGAA	1680
3093							
3094	GCAAATAAGA	TCTCACTGAA	GATTGGTAAA	GTCCAGCAAG	AAAATGAGAT	AAAAGAGAAG	1740
3095							
3096	CCTGAAATGA	CGAAAAAATC	AGGTGTTTTG	ATTCTTGGTG	CTGGACGTGT	GTGTCGCCCA	1800
3097							
3098	GCTGCTGATT	TCCTAGCTTC	AGTTAGAACC	ATTTCGTCAC	AGCAATGGTA	CAAAACATAT	1860
3099							
3100	TTCGGAGCAG	ACTCTGAAGA	GAAAACAGAT	GTTCATGTGA	TTGTCGCGTC	TCTGTATCTT	1920
3101							
3102	AAGGATGCCA	AAGAGACGGT	TGAAGGTATT	TCAGATGTAG	AAGCAGTTCG	GCTAGATGTA	1980
3103							
3104	TCTGATAGTG	AAAGTCTCCT	TAAGTATGTT	TCTCAGGTTG	ATGTTGTCCT	AAGTTTATTA	2040
3105							
3106	CCTGCAAGTT	GTCATGCTGT	TGTAGCAAAG	ACATGCATTG	AGCTGAAGAA	GCATCTCGTC	2100
3107							
3108	ACTGCTAGCT	ATGTTGATGA	TGAAACGTCC	ATGTTACATG	AGAAGGCTAA	GAGTGCTGGG	2160
3109							
3110	ATAACGATTC	TAGGCGAAAT	GGGACTGGAC	CCTGGAATCG	ATCACATGAT	GGCGATGAAA	2220
3111							
3112	ATGATCAACG	ATGCTCATAT	CAAAAAAGGG	AAAGTGAAGT	CTTTTACCTC	TTATTGTGGA	2280
3113							
3114	GGGCTTCCCT	CTCCTGCTGC	AGCAAATAAT	CCATTAGCAT	ATAAATTTAG	CTGGAACCCCT	2340
3115							
3116	GCTGGAGCAA	TTCGAGCTGG	TCAAAACCCC	GCCAAATACA	AAAGCAACGG	CGACATAATA	2400
3117							
3118	CATGTTGATG	GGAAGAATCT	CTATGATTCC	GCGGCAAGAT	TCCGAGTACC	TAATCTTCCA	2460
3119							
3120	GCTTTTGAT	TGGAGTGTTT	TCCAAATCGT	GA CTCCTTGG	TTTACGGGGA	ACATTATGGC	2520
3121							
3122	ATCGAGAGCG	AAGCAACAAC	GATATTTTCGT	GGAACACTCA	GATATGAAGG	GTTTAGTATG	2580
3123							
3124	ATAATGGCAA	CACTTTCGAA	ACTTGGATTG	TTTGACAGTG	AAGCAAATCA	AGTACTCTCC	2640
3125							
3126	ACTGGAAGA	GGATTACGTT	TGGTGCTCTT	TTAAGTAACA	TTCTAAATAA	GGATGCAGAC	2700
3127							
3128	AATGAATCAG	AGCCCCTAGC	GGGAGAAGAA	GAGATAAGCA	AGAGAATTAT	CAAGCTTGGA	2760
3129							
3130	CATTCCAAGG	AGACTGCAGC	CAAAGCTGCC	AAAACAATTG	TATTCTTGGG	GTTCAACGAA	2820
3131							
3132	GAGAGGGAGG	TTCCATCACT	GTGTAAAAGC	GTATTTGATG	CAACTTGTTA	CCTAATGGAA	2880
3133							
3134	GAGAACTAG	CTTATTCCGG	AAATGAACAG	GACATGGTGC	TTTTGCATCA	CGAAGTAGAA	2940
3135							
3136	GTGGAATTCC	TTGAAAGCAA	ACGTATAGAG	AAGCACACTG	CGACTCTTTT	GGAATTCGGG	3000
3137							
3138	GACATCAAGA	ATGGACAAAC	AACAACCGCT	ATGGCCAAGA	CTGTTGGGAT	CCCTGCAGCC	3060
3139							
3140	ATTGGAGCTC	TGGTGTTAAT	TGAAGACAAG	ATCAAGACAA	GAGGAGTCTT	AAGGCCTCTC	3120
3141							
3142	GAAGCAGAGG	TGTATTTGCC	AGCTTTGGAT	ATATTGCAAG	CATATGGTAT	AAAGCTGATG	3180

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/049,304DATE: 02/02/2000
TIME: 22:09:37

INPUT SET: S34610.raw

3143

3144 GAGAAGGCAG AATGA

3195

3145

3359 (2) INFORMATION FOR SEQ ID NO:113:

3360

3361 (i) SEQUENCE CHARACTERISTICS:

--> 3362 (A) LENGTH: 23 base pairs

3363 (B) TYPE: nucleic acid

3364 (C) STRANDEDNESS: single

3365 (D) TOPOLOGY: linear

3366

3367 (ii) MOLECULE TYPE: DNA (genomic)

3368

3369 (xi) SEQUENCE DESCRIPTION: 'SEQ ID NO:113:

3370

--> 3371 TTYTCTCAYACICAYAAAGCICA

23

3372

I is invalid for use in the sequence - use N and explain in

*(ix) FEATURE:
section*

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/049,304DATE: 02/02/2000
TIME: 22:09:37**INPUT SET: S34610.raw**

Line	Error	Original Text
325	# of Sequences for line conflicts w/ running total	CTT GAG GCT CTC CGA GAA GAC ATG AAA AAA GC
2023	# of Sequences for line conflicts w/ running total	AAG GCC ATG GAA GAG AAG ATG AAG GCG TGATA
3029	Entered (3194) and Calc. Seq. Length (3195) differ	(A)LENGTH: 3194 base pairs
3362	Entered (23) and Calc. Seq. Length (20) differ	(A)LENGTH: 23 base pairs
3371	Wrong Nucleic Acid Designator	TTYTCICAYA CICAYAARGC ICA
3371	Wrong Nucleic Acid Designator	TTYTCICAYA CICAYAARGC ICA
3371	Wrong Nucleic Acid Designator	TTYTCICAYA CICAYAARGC ICA
3371	# of Sequences for line conflicts w/ running total	TTYTCICAYA CICAYAARGC ICA